

We claim:

1. A method for inferring a network relationship between genes, comprising:

5 (a) providing a quantitative time course data library for a set of genes of an organism, said library including expression results based on time course of expression of each gene in said set of genes, quantifying an average effect and measure of variability of each time point on each other of said genes;

(b) creating a sparse matrix from said library, said matrix having zero coefficients removed therefrom;

10 (c) generating a set of linear differential equations from said matrix; and

(d) solving said set of equations to produce said network relationship.

2. The method of claim 1, wherein said zero coefficients are identified using Akaike's Information Criterion (AIC).

15

3. The method of claim 1, wherein said differential equation is

$$\frac{d}{dt}x(t) = \underline{\Lambda} \cdot x(t),$$

in which the vector $x(t)$ contains the amount of expressed cDNA as a function of time, and the matrix $\underline{\Lambda}$ is a constant with units second⁻¹.

20

4. The method of claim 1, wherein said matrix contains elements Λ_{ij} , wherein Λ_{ij} represents the effect of gene j on gene i , and wherein $[\Lambda_{ij}]^{-1}$ represents the reaction time for said effect of gene j on gene i .

25

5. The method of claim 1, wherein said differential equation solved is

$$\underline{x}(t) = \exp\left(\underline{\Lambda} t\right) \bullet \underline{x}_0,$$

6. The method of claim 1, wherein said exponent $\underline{\Lambda} t$ ($\exp(\underline{\Lambda})$) is solved using the formula:

$$\exp\left(\underline{\Lambda}\right) \equiv \sum_{i=0}^{\infty} \frac{1}{i!} \underline{\Lambda}^i.$$

5

7. The method of claim 1, wherein said differential equation is estimated by solving the difference equation:

$$\frac{\Delta x}{\Delta t} = \underline{\Lambda} \bullet x,$$

10 8. The method of claim 1, wherein said sparse matrix further comprises an error estimated using the formula:

$$\underline{x}(t + \Delta t) - \underline{x}(t) = \Delta t \bullet \underline{\Lambda} \bullet \underline{x}(t) + \underline{\varepsilon}(t).$$

15 9. The method of claim 8, wherein said error has a normal distribution independent of time according to the formula:

$$f(\underline{\varepsilon}(t); \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^m \exp\left\{-\frac{\underline{\varepsilon}(t)^T \bullet \underline{\varepsilon}(t)}{2\sigma^2}\right\},$$

wherein standard deviation σ is equal for each of said genes at all times.

20 10. The method of claim 1, wherein the maximum likelihood estimate of the variance σ^2 is determined by maximizing the log-likelihood function with respect to σ^2 using the formula:

$$\hat{\sigma}^2 = \frac{1}{n \cdot m} \sum_{i=1}^1 \hat{\underline{\underline{\varepsilon}}}_i^T \cdot \hat{\underline{\underline{\varepsilon}}}_i \cdot$$

11. The method of claim 10, wherein said variance σ^2 is determined using the formula:

$$\hat{\sigma}^2 = \frac{1}{nm} \sum_{i=1}^n \left[(\underline{x}_i^T - \underline{x}_{i-1}^T) \cdot (\underline{x}_i - \underline{x}_{i-1}) + (t_i - t_{i-1})^2 \underline{x}_{i-1}^T \cdot \underline{\underline{\Lambda}}^T \cdot \underline{\underline{\Lambda}} \cdot \underline{x}_{i-1} - 2(\underline{x}_i^T - (t_i - t_{i-1}) \underline{x}_{i-1}^T) \cdot \underline{\underline{\Lambda}} \cdot \underline{x}_{i-1} \right],$$

5

12. The method of claim 2, wherein said AIC is minimized using the formula:

$$AIC = 2 \cdot \left[\begin{array}{c} \text{log-likelihood of the} \\ \text{estimated model} \end{array} \right] + 2 \cdot \left[\begin{array}{c} \text{number of estimated} \\ \text{parameters} \end{array} \right] \quad (16)$$

13. The method of claim 1, wherein mask $\underline{\underline{M}}$ is used to set matrix elements of $\hat{\underline{\underline{\Lambda}}}$

10 equal to zero using the formula:

$$\hat{\underline{\underline{\Lambda}}} = \underline{\underline{M}} \circ \hat{\underline{\underline{\Lambda}}},$$

where \circ denotes an element-wise product, and mask $\underline{\underline{M}}$ is a matrix whose elements

are either one or zero.

15 14. The method of claim 13, wherein matrix elements are set to zero by applying a mask $\underline{\underline{M}}$ produced by minimizing the formula:

$$\text{if } M_{ij}=1: \left[\hat{\underline{\Lambda}}' \cdot \underline{A} \right]_{ij} = B_{ij};$$

$$\text{if } M_{ij}=0: \hat{\Lambda}_{ij}'=0,$$

thereby yielding the maximum likelihood estimate $\hat{\underline{\Lambda}}'$.

15. The method of claim 2, wherein said AIC is minimized according to the
5 formula:

$$AIC = n m \ln \left[2\pi \hat{\sigma}^2 \right] + nm + 2 \cdot (1 + [\text{sum of the maskelements } M_{ij}]) ,$$

16. The method of claim 13, wherein said mask \underline{M} is selected to minimize AIC
calculated using the formula:

$$AIC = n m \ln \left[2\pi \hat{\sigma}^2 \right] + nm + 2 \cdot (1 + [\text{sum of the maskelements } M_{ij}]) ,$$

10

17. A medium containing one or more results of network relationships between
genes calculated using a method of claim 1 stored thereon.

- 15 18. A method for determining the statistical significance of network relationships,
comprising:

- (a) calculating the average log-ratio of expression for each gene at each
time point;
- (b) calculating the standard deviation from all measurements;
- 20 (c) calculate the joint probability; and
- (d) adopting a criterion for statistical significance.

19. The method of claim 18, wherein said step (a) is determined using the formula:

$$\bar{x}_{ji} = \frac{1}{2} \sum_{k=1,2} x_{ji}[k].$$

5 20. The method of claim 18, wherein step (b) is determined using the formula:

$$\hat{\sigma}_{j|H_0} = \sqrt{\frac{1}{2n} \sum_{i=1}^n \sum_{k=1,2} (x_{ji}[k])^2},$$

in which $x_{ji}[k]$ is the data value of measurement k at time point i for gene j .

10 21. The method of claim 18, wherein the joint probability for $\bar{x}_{j\bullet}$ to be larger in absolute value than the measured values \bar{x}_{ji} is calculated using the formula:

$$\begin{aligned} P &= \prod_{i=1}^n P_i = \prod_{i=1}^n P(|\bar{x}_{j\bullet}| > |\bar{x}_{ji}|) \\ &= \prod_{i=1}^n \left[1 - \operatorname{erf}\left(\frac{|\bar{x}_{ji}|}{\hat{\sigma}_{j|H_0}/\sqrt{2}}\right) \right], \end{aligned}$$

15 wherein erf is an error function.

22. The method of claim 18, wherein a significance level α is selected.

23. The method of claim 18, wherein the null hypothesis is rejected if $P_i < \alpha$.

20

24. The method of claim 18, wherein the null hypothesis is rejected if $P < \alpha^n$, wherein n is the number of time points at which gene expression is evaluated.

25. A method for determining the statistical significance of network relationships, comprising:

(a) calculating the average log-ratio of measurements of expression for each gene at each time point using the formula:

$$\bar{x}_{ji} = \frac{1}{2} \sum_{k=1,2} x_{ji}[k].$$

(b) calculating the standard deviation of said measurements using the formula :

$$\hat{\sigma}_{j|H_0} = \sqrt{\frac{1}{2n} \sum_{i=1}^n \sum_{k=1,2} (x_{ji}[k])^2},$$

in which $x_{ji}[k]$ is the data value of measurement k at time point i for gene j .

(c) calculating a joint probability for $\bar{x}_{j\bullet}$ to be larger in absolute value than measured values \bar{x}_{ji} calculated using the formula:

$$\begin{aligned} P &= \prod_{i=1}^n P_i = \prod_{i=1}^n P(|\bar{x}_{j\bullet}| > |\bar{x}_{ji}|) \\ &= \prod_{i=1}^n \left[1 - \operatorname{erf}\left(\frac{|\bar{x}_{ji}|}{\hat{\sigma}_{j|H_0}/\sqrt{2}}\right) \right], \end{aligned}$$

wherein erf is an error function; and

(d) applying a criterion for statistical significance to determine whether a null hypothesis is rejected.

- 5 26. The method of claim 25, wherein the null hypothesis is rejected if $P < \alpha^n$, wherein n is the number of time points at which gene expression is evaluated.